Disclosed herein are methods and systems for managing information management system I/O resources (*e.g.,* file system resources, storage system resources, *etc.*) in information delivery environments. The disclosed methods and systems may be configured to employ unique resource modeling and/or resource monitoring techniques and may be advantageously implemented in a variety of information delivery environments and/or with a variety of types of information management systems. Examples of just a few of the many types of information delivery environments and/or information management system configurations with which the disclosed methods and systems may be advantageously employed are described in co-pending United States patent application serial number 09/797,413 filed on March 1, 2001 which is entitled NETWORK CONNECTED COMPUTING SYSTEM; in co-pending United States patent application serial number 09/797,200 filed on March 1, 2001 which is entitled SYSTEMS AND METHODS FOR THE DETERMINISTIC MANAGEMENT OF INFORMATION; and in co-pending United States patent application serial number 09/879,810 filed on June 12, 2001 which is entitled SYSTEMS AND METHODS FOR PROVIDING DIFFERENTIATED SERVICE IN INFORMATION MANAGEMENT ENVIRONMENTS; each of the foregoing applications being incorporated herein by reference.

Included among the examples of information management systems with which the disclosed methods and systems may be implemented are network content delivery systems that deliver non-continuous content (*e.g.,* HTTP, *etc.*), continuous streaming content (e.g., streaming video, streaming audio, web proxy cache for Internet streaming, etc.) and/or that deliver over-size or very large data objects of any other kind, such as over-size non-continuous data objects. As used herein an "over-size data object" refers to a data object that has an object size that is so large relative to the available buffer/cache memory size of a given information management system, that caching of the entire data object is not possible or is not allowed by policy within the given system. Examples of non-continuous over-size data objects include, but are not limited to, relatively large FTP files, *etc.*

By monitoring resource consumption and availability (*e.g.,* disk workload, logical volume workload, disk capacity, *etc.*) resource characteristics may be modeled and used to make admission control decisions and to define read-ahead strategy. Dynamic resource monitoring may be further implemented to enable dynamic and/or adaptive I/O resource management, for example, to make admission control decisions and/or adjust read-ahead strategy as desired or needed based on changing characteristics of resource consumption/availability characteristics. Such an adaptive approach to I/O resource modeling and management makes possible enhanced system I/O performance to fit a variety of changing information management system I/O conditions. In one exemplary embodiment, dynamic measurement-based I/O admission control may be enabled by monitoring the workload and the storage device utilization constantly during system run-time, and accepting or rejecting new I/O requests based on the run-time knowledge of the workload. In this regard, workload may be expressed herein in terms of outstanding I/O's or read requests.

The disclosed methods and systems may be implemented to manage memory units stored in any type of memory storage device or group of such devices suitable for providing storage and access to such memory units by, for example, a network, one or more processing engines or modules, storage and I/O subsystems in a file server, etc. Examples of suitable memory storage devices include, but are not limited to random access memory ("RAM"), disk storage, I/O subsystem, file system, operating system or combinations thereof. Memory units may be organized and referenced within a given memory storage device or group of such devices using any method suitable for organizing and managing memory units. For example, a memory identifier, such as a pointer or index, may be associated with a memory unit and "mapped" to the particular physical memory location in the storage device (*e.g.* first node of Q_1^{used} = location FF00 in physical memory). In such an embodiment, a memory identifier of a particular memory unit may be assigned/reassigned within and between various layer and queue locations without actually changing the physical location of the memory unit in the storage media or device. Further, memory units, or portions thereof, may be located in non-contiguous areas of the storage memory. However, it will be understood that in other embodiments memory management techniques that use contiguous areas of storage memory and/or that employ physical movement of memory units between locations in a storage device or group of such devices may also be employed.

Partitioned groups of storage devices may be present, for example, in embodiments where resources (*e.g.*, multiple storage devices, buffer memory, *etc.*) are partitioned into groups on the basis of one or more characteristics of the resources (*e.g.*, on basis of physical drives, on basis of logical volume, on basis of multiple tenants, *etc.*). In one such embodiment, storage device resources may be associated with buffer memory and/or other resources of a given resource group according to a particular resource characteristic, such as one or more of those characteristics just described.

Although described herein in relation to block level memory, it will be understood that embodiments of the disclosed methods and system may be implemented to manage memory units on virtually any memory level scale including, but not limited to, file level units, bytes, bits, sector, segment of a file, etc. However, management of memory on a block level basis instead of a file level basis may present advantages for particular memory management applications, by reducing the computational complexity that may be incurred when manipulating relatively large files and files of varying size. In addition, block level management may facilitate a more uniform approach to the simultaneous management of files of differing type such as HTTP/FTP and video streaming files.

The disclosed methods and systems may be implemented in combination with any memory management method, system or structure suitable for logically or physically organizing and/or managing memory, including integrated logical memory management structures such as those described in United States Patent Application Serial No. 09/797,198 filed on March 1, 2001 which is entitled SYSTEMS AND METHODS FOR MANAGEMENT OF MEMORY; and in United States Patent Application Serial No. 09/797,201 filed on March 1, 2001 which is entitled SYSTEMS AND METHODS FOR MANAGEMENT OF MEMORY IN INFORMATION DELIVERY ENVIRONMENTS, each of which is incorporated herein by reference. Such integrated logical memory management structures may include, for example, at least two layers of a configurable number of multiple memory queues (e.g., at least one buffer layer and at least one cache layer), and may also employ a multi-dimensional positioning algorithm for memory units in the memory that may be used to reflect the relative priorities of a memory unit in the memory, for example, in terms of both recency and frequency. Memory-